



**HAL**  
open science

## Extração de Propriedades Sociais em Redes Veiculares

Felipe Domingos da Cunha, Aline Carneiro Viana, Thiago Rodrigues, Raquel A. F. Mini, Antonio A. F. Loureiro

► **To cite this version:**

Felipe Domingos da Cunha, Aline Carneiro Viana, Thiago Rodrigues, Raquel A. F. Mini, Antonio A. F. Loureiro. Extração de Propriedades Sociais em Redes Veiculares. XXXII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos, SBC, May 2014, Florianópolis, Brazil. hal-00981189

**HAL Id: hal-00981189**

**<https://inria.hal.science/hal-00981189>**

Submitted on 21 Apr 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Extração de Propriedades Sociais em Redes Veiculares

Felipe Cunha<sup>1,3</sup>, Aline Viana<sup>3</sup>, Thiago Rodrigues<sup>2</sup>, Raquel Mini<sup>2</sup>, Antonio Loureiro<sup>1</sup>

<sup>1</sup>DCC – Universidade Federal de Minas Gerais

<sup>2</sup>DCC – Pontifícia Universidade Católica de Minas Gerais

<sup>3</sup>INRIA - Saclay - France

{fdcunha, loureiro}@dcc.ufmg.br, aline.viana@inria.fr  
taorodrigues@sga.pucminas.br, raquelmini@pucminas.br

**Abstract.** *Everyday, vehicles perform routes in a city and along of their trajectories they encounter other vehicles. The frequency of these encounters is influenced by many factors, such as: vehicle speed, destinations, traffic conditions, and the period of the day. However, these factors are justified by the public roads limits and the driver's behavior. The people present daily routines and similar behaviors that have a great impact in the daily traffic evolution. In this work, we present a numerical analysis of real and realistic traces data set that describe the mobility of a set of vehicles. Social metrics are computed, and the results obtained are compared to random graphs in the direction to attest that Vehicular Network presents a social behavior. Results show the existence of routines and social properties in Vehicular Networks. Finally, we discuss about new social perspectives in Vehicular Networks.*

**Resumo.** *Todos os dias, veículos transitam nas cidades e durante suas trajetórias encontram outros veículos. A frequência destes encontros pode ser influenciada por diversos fatores como: a velocidade dos veículos, os destinos escolhidos, as condições de tráfego e o período do dia. Entretanto, estes fatores são justificados pelos limites das vias públicas e pelo comportamento dos motoristas. Os comportamentos similares dos motoristas e suas rotinas diárias têm um grande impacto na evolução do tráfego diariamente. Neste trabalho, é apresentada uma análise numérica de traces reais e realísticos que descrevem a mobilidade de um conjunto de veículos. Para isso, métricas sociais são computadas e os resultados obtidos são comparados com grafos aleatórios na direção de atestar que as Redes Veiculares apresentam comportamento social. Os resultados comprovam a existência de rotinas e propriedades sociais nas Redes Veiculares. Finalizando, discute-se sobre novas perspectivas sociais em Redes Veiculares.*

## 1. Introdução

As Redes Veiculares (*Vehicular Networks - VANETs*) são um tipo especial de redes ad-hoc, formadas por veículos que possuem capacidade de processamento e comunicação sem fio, transitando por ruas e rodovias. Geralmente, esses veículos podem se comunicar diretamente ou pelo uso de pontos de acesso (RSUs). Através das RSUs pode-se ter acesso e obter dados de outras redes, como por exemplo a Internet. Devido a esta natureza, as redes veiculares podem estabelecer a comunicação em diferentes ambientes, como centros urbanos e rodovias [Boukerche et al. 2008].

Essas redes são propostas para ajudar motoristas com o trânsito, para evitar colisões com a emissão de alertas, para prover entretenimento para motoristas e passageiros, e também para criar um sistema de tráfego inteligente. Uma importante aplicação é a emissão de alertas para os motoristas, com informações acerca das condições das rodovias e perigos de colisões iminentes. Estudos recentes mostram que 60% dos acidentes de veículos podem ser evitados se os condutores forem avisados segundos antes da colisão [Faezipour et al. 2012]. Assim, uma interação entre os veículos pode contribuir fortemente para a disseminação de uma mensagem de alerta, e consequentemente reduzir o número de acidentes [Schoch et al. 2008]. Outro exemplo de aplicações é o auxílio aos condutores, com o objetivo de fornecer um melhor caminho. Com isso, é importante que essas redes monitorem e prevejam as condições de tráfego com sugestões de mudanças de rota, para evitar engarrafamentos e assim, reduzir o tempo da viagem [Tostes et al. 2013].

Nestas redes a comunicação é altamente influenciada pela densidade de veículos e os padrões de mobilidade dos mesmos. Veículos se movem e param de acordo com os limites de velocidade e direção impostos pelas vias públicas. Entretanto, o período do dia também pode influenciar na mobilidade. Em horários de pico, o tráfego pode ser mais lento, com engarrafamentos; o que força o motorista a parar o veículo ou mudar a sua rota. Em contrapartida, os finais de semana, feriados e madrugadas são períodos nos quais o tráfego é ocioso. Esta variação de densidade nos horários reforça a característica dinâmica das redes veiculares, tornando a comunicação uma tarefa desafiadora.

A mobilidade também é influenciada pelo comportamento do motorista e sua rotina. Durante os finais de semana, destinos como: sítio, fazenda, igreja e centros comerciais são escolhidos para lazer e entretenimento, e são frequentemente visitados. Nos dias de semana, um comportamento contrário é observado. Pessoas repetem seus caminhos diariamente, nos mesmos períodos, para os mesmos destinos como: escolas, trabalhos, universidades, restaurantes, etc. Durante suas trajetórias encontram com outros veículos, atravessam as mesmas vias e estão sujeitas as mesmas condições de tráfego. Estas características sugerem o estudo da mobilidade veicular sob uma perspectiva social, com o intuito de entender melhor sua evolução e de aplicar os conceitos sociais obtidos para melhorar as aplicações e a conectividade em redes veiculares.

Nesta direção, o objetivo deste trabalho é o de investigar a seguinte questão: *Quais são as características que uma rede veicular precisa apresentar para comprovar existência de um comportamento social?* Assim, este trabalho apresenta uma análise social de duas bases de dados (*traces*) que têm características de diferentes grupos de veículos. A primeira base é realística e contém registros de dados de veículos durante uma rotina diária. A segunda base é real e contém registros da mobilidade de táxis durante um mês. Esta análise será efetuada a partir do registro de movimento dos veículos e seus encontros, modelados como grafos temporais. Nestes grafos serão computadas métricas que certifiquem a existência de propriedades sociais. Para melhorar a compreensão dos resultados e tornar a análise consistente, algumas métricas serão calculadas também em grafos aleatórios, que serão gerados com a mesma quantidade de vértices e distribuição de graus dos grafos temporais. Finalizando, serão apresentados cenários e situações nos quais a aplicação das métricas sociais em redes veiculares é relevante.

O restante do trabalho está organizado da seguinte forma. A Seção 2 apresenta um breve resumo dos trabalhos relacionados relevantes da literatura. A Seção 3 descreve

a metodologia utilizada para avaliação das bases de dados. A Seção 4 apresenta as bases de dados utilizadas e os resultados quantitativos, de acordo com cada métrica avaliada. Completando, a Seção 5 expõe as conclusões finais e perspectivas futuras.

## 2. Trabalhos Relacionados

“As redes de computadores são inerentemente redes sociais, ligando as pessoas, organizações e conhecimento” [Wellman 2001]. Nesta perspectiva, o conceito de redes sociais tem sido explorado em diferentes tipos de redes de computadores, principalmente para melhor entender a evolução da topologia e para analisar os encontros físicos dos nós, com o objetivo de melhorar o desempenho dos protocolos de comunicação [Scott 2000, Hossmann et al. 2009].

Entretanto, em redes veiculares, as análises sociais podem ser usadas para monitorar a evolução do tráfego durante o dia, objetivando o conhecimento das rotinas dos usuários, identificando trajetórias similares e horários de pico. O objetivo dessas análises é melhorar e adaptar os serviços providos neste tipo de rede. Considerando o comportamento social, em [Fiore and Härrri 2008], Fiore et al. apresentam uma análise detalhada das propriedades topológicas de uma rede veicular. Os autores usam métricas sociais para investigar a evolução temporal da topologia da rede. Resultados mostraram que é possível tirar vantagem da mobilidade veicular para melhorar o desempenho dos protocolos de rede. Contudo, os autores efetuam a análise se baseando apenas em modelos de tráfego analíticos, conhecidos na literatura.

Neste mesmo contexto, em [Loulloudes et al. 2010], os autores discutem como as métricas sociais podem ser empregadas para melhorar o desempenho dos protocolos de roteamento em redes veiculares. A análise, entretanto, é baseada em um curto período de 2 horas. Acredita-se que um curto período de observação não é suficiente para extrair comportamentos sociais de uma mobilidade veicular. Outra análise social é apresentada em [Liu et al. 2012]. Este trabalho discute diversas leis universais das redes sociais e apresenta uma análise numérica de uma real base de dados. Os autores escolheram duas bases públicas: São Francisco e Shanghai, que descrevem os movimentos de táxis. As métricas sociais foram estimadas e os resultados mostraram a existência de rotinas sociais em VANETs, comprovando as leis de potência para a distribuição de graus e o fenômeno *small world*. Este trabalho considera apenas a mobilidade da base de dados de táxis e avalia somente as métricas sociais, sem verificar a existência de rotinas.

Em [Uppoor and Fiore 2012] também é apresentado uma análise de base de dados de veículos. Entretanto os autores se atêm apenas no estudo e classificação das trajetórias diárias dos condutores. A base de dados avaliada foi da cidade de Cologne e os autores a caracterizaram de forma macro e microscópica. Como macroscópica, com a análise da evolução da densidade de o dinamismo dos fluxos de veículos entre as regiões na cidade. E microscópica, com a análise da distribuição de veículos por área e a quantidade de encontro entre eles. Contudo, esta análise não considera aspectos sociais e tampouco a interação entre os veículos.

Como estudo inicial, em [Cunha et al. 2013], apresentamos uma avaliação de métricas sociais em base de dados de veículos com duração de 24 horas. Esta avaliação apresentou a existência de rotinas sociais e comportamentos similares em redes veiculares. Porém, este estende e aprimora o estudo com a análise de bases de dados com longa

duração. Os resultados obtidos com esta avaliação serão comparados com resultados das métricas aplicadas a grafos aleatórios. Além disso, também será indicado possíveis serviços e protocolos das redes veiculares que poderão fazer uso destas métricas para alcançar melhorias de desempenho.

### 3. Metodologia

Nesta seção serão apresentados os detalhes da metodologia usada em toda a avaliação das bases de dados. A execução desta avaliação consiste na construção de grafos que representam as interações entre os veículos e o cálculo das métricas sociais. Assim, nas seções a seguir serão descritas as suposições feitas na construção dos grafos temporais e randômicos e a definição das métricas avaliadas.

#### 3.1. Grafos Temporais

O modelo utilizado para construção dos grafos temporais considera a existência de duas arestas entre os vértices quando dois veículos tiveram um encontro, isto é, um veículo está localizado fisicamente dentro do raio e comunicação do outro. Nessa avaliação o raio de comunicação considerado foi de 100 metros, de acordo com o protocolo 802.11p e o cenário de centros urbanos. Conforme padronização entre as bases avaliadas, a velocidade média dos veículos foi de 25 km/h. Assim, de acordo com esta velocidade e o raio de comunicação, os registros de posição dos veículos são analisados. Sempre que houver dois veículos próximos e a diferença do registro de data/hora entre eles for de 15 s., um encontro será registrado no grafo.

O grafo então é definido considerando os encontros que aconteceram hora-a-hora. O grafo no instante  $t$  é unidirecional e pode ser definido como um grafo  $G(t) = (V, E)$ , onde  $V$  representa o conjunto de veículos com todos veículos  $v_i$  e  $E$  representa o conjunto de arestas. O grafo  $G$  é temporal, no qual uma aresta  $e_{ij}(t)$  existe entre dois veículos  $v_i$  e  $v_j$  somente durante o tempo  $t$ , com  $i \neq j$ . Todas as métricas foram avaliadas de hora em hora, considerando o grafo tempo  $G(t)$ .

#### 3.2. Grafos Aleatórios

Um grafo aleatório é um grafo obtido a partir de um conjunto de vértices  $V$  e sua respectiva distribuição de graus. A adição de arestas no grafo  $G^R$  é feita de forma aleatória, baseado em modelos de probabilidade. Neste trabalho, o processo de construção dos grafos aleatórios  $G^R$  baseia-se no algoritmo de urna, usado na geração de estruturas aleatórias [Johnson and Kotz 1977]. Nele, o primeiro passo consiste na inserção de  $d_i$  “bolas” em uma urna, marcadas com o identificador  $i$  do vértice, de forma que  $d_i$  corresponde ao grau do vértice  $i$ . Feito isso para todos os vértices, duas bolas  $b_i$  e  $b_j$  são retiradas de forma aleatória e seus identificadores são verificados. Se  $i \neq j$  e não existe aresta entre esses dois vértices, uma aresta então será inserida. Esse processo é repetido, até que não exista nenhuma bola na urna ou não seja possível conectar os vértices presentes na urna. Assim, quando não for possível conectar os vértices, haverá uma diferença entre o número de arestas do grafo original  $G$  e  $G^R$ . A diferença permitida pelo processo poderá ser de até 0.001%, que não compromete a avaliação. A partir deste algoritmo foram gerados grafos aleatórios  $G^R$  para cada grafo temporal  $G(t)$  analisado, nos quais as métricas também serão computadas.

### 3.3. Métricas

As métricas sociais escolhidas para avaliação serão descritas, considerando os aspectos e características da mobilidade dos veículos. Quanto a mobilidade dos veículos, pode-se classificar as características em macroscópicas e microscópicas [Uppoor and Fiore 2012]. Nesse contexto, o trabalho estende essa classificação para as métricas sociais, dividindo-as de acordo com o que cada uma representa no grafo. As métricas macroscópicas representam medidas do estado global da rede, as quais podem retratar um comportamento geral de todos os veículos e da evolução do grafo temporal. As métricas microscópicas definem valores individuais os veículos, representando o comportamento de um único veículo. Para as métricas macroscópicas foram escolhidas a Distância, a Densidade, o Diâmetro e a Persistência das Arestas. Como métricas microscópicas foram selecionadas o Grau do Vértice, o Coeficiente de Agrupamento e Centralidade de Proximidade. Todas as métricas escolhidas são conhecidas na literatura e comumente utilizadas em trabalhos de avaliação de aspectos sociais em grafos.

#### 3.3.1. Métricas Macroscópicas

*Distância:* esta é uma métrica que compreende o tamanho do caminho entre um par de veículos, em termos do número de saltos necessários. Em redes veiculares, a distância pode representar a existência de um interesse comum entre condutores. Se a distância entre os veículos  $v_i$  e  $v_j$  é grande, isto implica em dizer que eles estão fisicamente distantes, fora do raio de alcance, e provavelmente não visitam lugares próximos. Caso contrário, se a distância é pequena,  $v_i$  e  $v_j$  podem visitar lugares próximos ou possuem pontos em comum ao longo de sua rota. Geralmente, grafos com o comportamento de *small world* possuem curta distância.

*Diâmetro:* pode-se definir o diâmetro de um grafo como a maior distância entre quaisquer pares de nós. Considerando as redes veiculares, um grafo com muitos veículos e um pequeno diâmetro, pode indicar existência de pontos de interesse em comum entre esses veículos ou também a possibilidade de engarrafamentos. Em contrapartida, um grande diâmetro pode indicar a não existência de similaridade na rotina dos motoristas. Além disso, um diâmetro maior pode acarretar o custo na comunicação, com mais atrasos entre a origem e o destino, devido à quantidade de saltos que a informação fará.

*Densidade:* representa o quão densa é a rede, isto é, a quantidade de conexões existentes entre os veículos. No contexto das redes veiculares, regiões urbanas podem ter uma alta densidade que regiões periféricas ou rurais. Em grandes centros urbanos, a densidade tende a ser elevada devido ao tráfego lento e a existência de engarrafamentos. Esta métrica pode ser computada pela razão entre o número de arestas existentes no grafo  $e(t)$  e o número de arestas de um grafo completo  $|V||V - 1|/2$ .

*Persistência de Arestas:* é uma métrica que representa a persistência de um encontro entre dois veículos. Esta análise considera persistência de arestas o número de vezes que o veículo  $v_i$  e  $v_j$  se encontram durante o tempo. Para isso, computa-se essa métrica em função do peso da aresta que representa o número de vezes que o encontro  $e_{ij}(t)$  acontece.

### 3.3.2. Métricas Microscópicas

*Grau do Vértice*: determinar o número de encontros distintos que um veículo teve durante um período. Esta métrica pode ser influenciada pela trajetória do veículo e pelo período do dia. Se um veículo cruza uma região de alta densidade ou em horário de pico, o seu grau tende a ser elevado. De outra forma, em regiões afastadas do centro da cidade ou em horas de baixo tráfego, o grau pode ser baixo e até nulo. Como uma definição formal, o grau de um veículo  $v_i$  no tempo  $t$  pode ser definido pela expressão:  $Grau_i(t) = ||\{v_j | \exists e_{ij}(t)\}||$ .

*Coefficiente de Agrupamento (cluster coefficient)*: esta métrica pode avaliar o quão próximo estão os vizinhos de um veículo  $v_i$  no grafo, isto é, a probabilidade que dois vizinhos do veículo  $v_i$  tem de se encontrem. Esta métrica pode representar regiões de engarrafamento, quando o coeficiente é alto. Além disso, se a rede apresenta o fenômeno *small world*, seu coeficiente de agrupamento deve ser elevado [Watts and Strogatz 1998].

*Centralidade de Proximidade (closeness centrality)*: esta métrica mensura a centralidade de um veículo em relação a sua distância para os demais veículos do grafo. O quão mais central for o veículo, menor será a sua distância para os demais. Apesar de ser uma métrica global, seu valor também pode representar a centralidade da rota de um veículo em relação aos demais. Possivelmente, veículos que trafegam mais em áreas centrais tendem a ter um valor baixo de centralidade.

## 4. Resultados

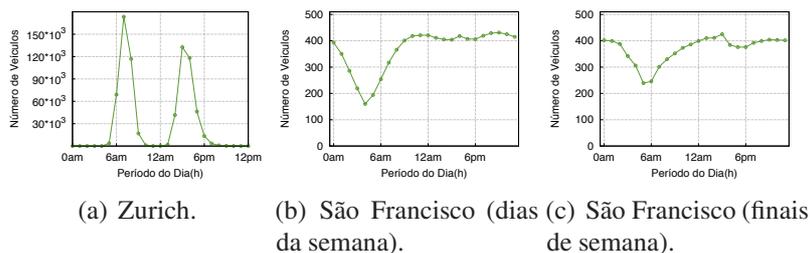
Esta seção apresenta os resultados da análise efetuada sob as bases de dados. A Seção 4.1 descreve as características de cada base de dados avaliada. A Seção 4.2 apresenta os resultados para as métricas macroscópicas e a Seção 4.3 apresenta os resultados para as métricas microscópicas de acordo com a definição na Seção 3.3.

### 4.1. Traces Avaliados

Foram analisadas duas bases de dados públicos (*traces*), que descrevem a mobilidade de veículos. A primeira base de dados é uma base realística, no qual os dados foram gerados baseados em modelos matemáticos que consideram informações de tráfego da cidade de Zurich [Naumov et al. 2006]. Esta base de dados representa a evolução do tráfego urbano durante um dia, incluindo a variação de horários de pico e madrugadas. São informações geradas por um simulador, e contém detalhes de destinos dos veículos e a velocidade empregada para efetuar o trajeto. Pode-se ver essa variação do tráfego na Figura 1-(a). Nota-se que existem dois intervalos de tempo os quais apresentam uma grande quantidade de veículos, chamados de horários de pico: 6 – 9am e 2 – 5pm. Toda a base de dados descreve o movimento de 260.000 veículos, com a concentração de veículos nos horários de pico: 173.470 veículos às 7am, e 132.633 veículos às 3pm. Essa base de dados foi analisada considerando todo o dia, com o destaque para os horários de pico, de acordo com cada métrica.

A segunda base de dados avaliada é uma base real, que contém o registro de mobilidade de 551 táxis na cidade de São Francisco (EUA) durante quatro semanas [Piorkowski et al. 2009]. São registros reais, coletados por dispositivos GPS localizados em cada veículo, minuto-a-minuto, durante sua trajetória. É possível observar a distribuição de veículos durante os dias da semana e finais de semana, respectivamente, nas Figuras 1-(b)

e 1-(c). Pode-se verificar nos dois gráficos um tráfego menos intenso durante a madrugada. Entretanto, durante os finais de semana o decaimento da quantidade de táxis em circulação é menor. Isso pode ser atribuído ao fato das pessoas permanecerem na rua até mais tarde, tendo mais eventos noturnos. Durante a semana é possível verificar o início das jornadas diárias de circulação dos veículos, com o crescimento da curva a partir das 6am. Além disso, percebe-se um comportamento constante na quantidade de veículos, com mais de 400 táxis em circulação durante todo o dia. Devido a esse comportamento, nesta base, as métricas foram calculadas considerando os cenários: dias da semana e finais de semana.



**Figura 1. Distribuição do Tráfego durante o dia.**

Nas seções a seguir serão apresentados os resultados numéricos para cada métrica avaliada em cada uma das bases de dados. Com o efeito de comparação, deve-se lembrar de que alguns resultados serão apresentados com o seu valor correspondente computado para grafos aleatórios. Esses grafos foram gerados com o mesmo número de vértices e a mesma distribuição de graus para cada período do dia.

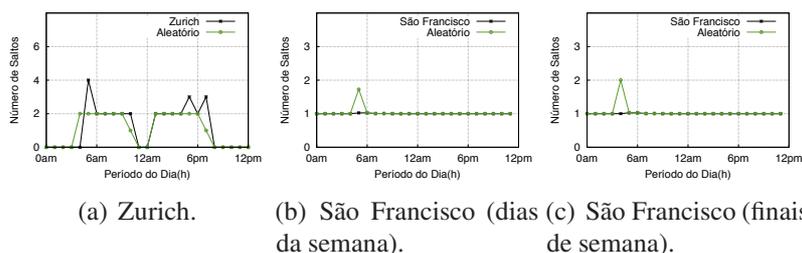
## 4.2. Métricas Macroscópicas

Nesta seção serão descritos os resultados para as métricas que retrata a estrutura global do grafo, e que no contexto veicular, conseguem capturar informações de como o comportamento de todos os veículos pode influenciar nessa estrutura e por consequência na mobilidade.

### Distância:

Ao analisar a distância entre dois vértices num grafo, é importante lembrar neste caso que a distância considera o número médio de arestas que conectam dois vértices, sem considerar o peso das arestas. Numa rede veicular, esta métrica pode representar se dois veículos frequentam ou não o mesmo lugar, isto é, veículos que possuem rotinas similares tendem a apresentar distância curta entre eles. Em compensação, veículos que transitam em horários diferentes, ou atravessam regiões distantes, tendem a possuir longas distâncias para os demais. O gráfico na Figura 2 apresenta a média da distância para as duas bases analisadas.

Na Figura 2-(a) verifica-se a distância na base de Zurich. Nos horários nos quais não existe tráfego a distância foi considerada 0. Entretanto, nos demais horários é possível observar o aumento do valor da distância quando há mais tráfego. Todavia, esse aumento não é tão expressivo, pois, apesar do tráfego intenso, existem pontos congruentes de rotas nas regiões centrais, o que tende a manter o valor da distância baixo, considerando que novos encontros entre os veículos acontecerão. Ao analisar o comportamento da curva



**Figura 2. Evolução da Distância durante o dia.**

para os grafos aleatórios, é possível verificar que uma distância menor em alguns horários. Devido à forma de geração, a inserção de algumas arestas pode ter criado atalhos no grafo, reduzindo a distância média entre os vértices. De diferente maneira, nas Figuras 2-(b) e 2-(c), pode-se observar o comportamento da distância para a base de São Francisco. Nota-se um pequeno aumento na distância no momento de pouco tráfego de táxis nas ruas. Esse comportamento é esperado pois, com a redução de veículos a tendência também é reduzir os encontros, o que reflete no valor da distância. Nos demais horários, a distância apresenta um comportamento constante, próximo de 1. Desta forma, a comparação com grafos aleatórios não apresentou variações. Para a base de Zurich pode-se constatar o fenômeno *small world*: um grande número de vértices no grafo com pequena distância entre eles. Entretanto, devido a estrutura do grafo, não é possível verificar o mesmo comportamento para a base de São Francisco.

### Diâmetro:

Em concordância com a métrica distância, o diâmetro do grafo consegue capturar a maior distância entre dois vértices do grafo. Desta forma, com essa métrica é possível verificar maiores detalhes da estrutura topológica do grafo e entender melhor a evolução da mesma. Importante lembrar que a avaliação dessa métrica só é possível se o grafo for conexo. Caso contrário, o diâmetro representará a maior distância entre dois vértices, dentre os componentes do grafo. Para as redes veiculares, um grafo  $G(t)$  com muitos vértices e um diâmetro pequeno pode indicar a ocorrência de engarrafamentos, ou mesmo a existência de uma grande similaridade de rotas entre os veículos para o instante analisado. Neste contexto, um grafo  $G(t)$  com poucos vértices e um grande diâmetro, pode indicar a existência de tráfego periférico com veículos possuindo rotas com poucos encontros e pouca similaridade de comportamento. Além disso, ao considerar o custo de comunicação, um grande diâmetro pode acarretar num aumento de custo para transmissão, sendo necessário mais saltos para a informação alcançar o destino.

A Figura 3 apresenta a evolução do valor do diâmetro durante o dia. Pode-se observar para a base de Zurich (Fig. 3-(a)) que o valor diâmetro sofre uma redução nos horários de pico. Isso ocorre, pois nestes horários os encontros nas regiões centrais são mais frequentes. Para a base de São Francisco (Fig. 3-(b) e 3-(c)) pode-se observar um comportamento semelhante com o aumento do diâmetro nos horários de redução de tráfego. Nesse período, com a redução do trânsito de veículos nas ruas, a probabilidade de encontro também reduz e a distância entre os mesmos se torna maior.

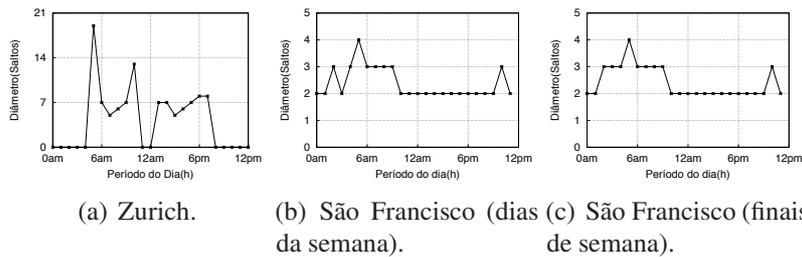


Figura 3. Evolução da Diâmetro durante o dia.

### Densidade:

Na análise desta métrica é avaliada a densidade global de um grafo. Ela representa os encontros que aconteceram durante um dia e a Figura 4 apresenta sua evolução. A primeira percepção possível nesta figura é quando se compara a evolução diária das duas bases. Táxis normalmente estão em circulação sempre e veículos particulares tem rotinas mais definidas. A Figura 4-(a) apresenta a densidade para a base de dados de Zurich. Pode-se constatar que, devido ao cenário, essa base apresenta uma densidade baixa. E seu valor não excede 2% durante todo o dia. Isto acontece pois, veículos estão sempre sujeitos a restrições de velocidade e direção das vias públicas e, neste contexto, podem estar próximos mas não acontecerão o contato. Outro fator que pode influenciar a existência de encontros é o engarrafamento. Este limita a velocidade do veículo e muitas vezes obriga o motorista a optar por atalhos, o que impede a existência de outros encontros.

Para o trace de São Francisco, a Figura 4-(b) e 4-(c) apresenta respectivamente a evolução para os dias da semana e finais de semana. Nota-se que durante o dia a densidade permanece superior a 80% e durante a madrugada esse valor reduz para 60%. Isso ocorre pois o número de viagens feitas por um táxi durante o dia é maior e também, por haver um maior número de táxis em circulação, a probabilidade de ocorrer encontros é maior. Em contrapartida, na madrugada, pela redução de táxis em circulação e a própria ociosidade de corridas neste horário, contribui para a redução da quantidade de encontros. Quando se compara os dois gráficos, nota-se uma pequena diferença na densidade entre os dias da semana e os finais de semana. Essa diferença pode ser atribuída à escolha de destinos nos finais de semana. Por serem destinos mais variados, a tendência é que ocorra uma redução na probabilidade de encontro entre os táxis.

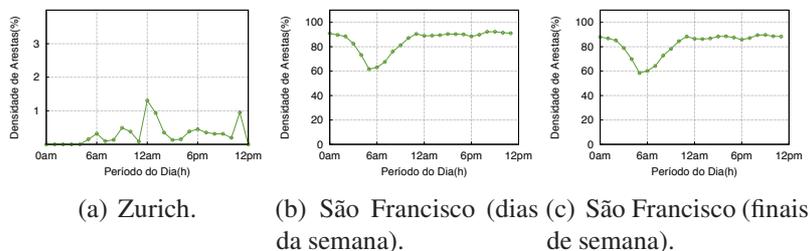


Figura 4. Evolução da Densidade durante o dia.

### Persistência de Arestas:

Neste trabalho, persistência de arestas descreve a quantidade de vezes que o par de arestas

tas de repete ao longo do tempo, ou seja, se um encontro entre dois veículos se repete. Entretanto, devido as diferentes durações das bases de dados, avaliamos a persistência de duas formas. No cenário de Zurich, a persistência é mensurada considerando os horários de pico, nos quais a concentração de veículos é maior. Desta forma foi possível capturar os encontros que acontecem no início do dia e que se repetem no final do dia. Para o trace de São Francisco, consideramos a persistência dia-a-dia, no mesmo horário. Com a avaliação dessa métrica, o foco é verificar nas duas bases de dados a porcentagem de veículos que possuem o comportamento similar, isto é, passam pelos mesmos locais e se encontram com frequência.

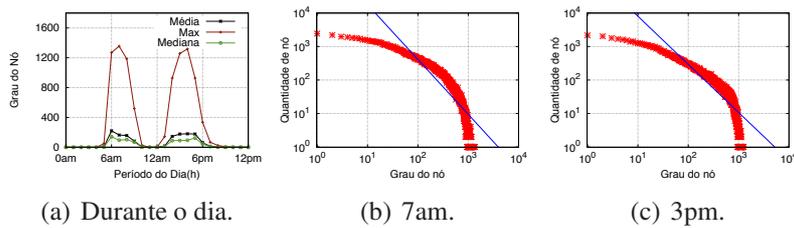
Para a base de dado de Zurich, no início do dia (6 – 9am.), o grafo possui 376.245 arestas e no final do dia (2 – 5pm.) o grafo possui 338.513 arestas. A persistência de arestas para esta base foi de 43% das arestas, isto é, aproximadamente 150.000 arestas se repetem, o que indica uma parte expressiva de veículos tem o comportamento semelhante na cidade, com a mesma rotina. Para a base dados de São Francisco também foi considerado os horários de maior quantidade de veículos em circulação, que começa a acontecer a partir das 6am. Nesse momento, a quantidade de veículos permanece superior a 400 táxis e a quantidade de arestas em média no grafo é de 60.000 arestas. Assim, com a porcentagem de 75%, a persistência de arestas para a base de São Francisco é maior. Apesar de serem veículos que trafegam para destinos diversos, táxis tendem a voltar para o seu ponto de partida, no qual aguardam novos passageiros. Além disso, esses pontos normalmente são fixos, o que aumenta a probabilidade de encontrar os mesmos veículos. Este é um bom indicativo para comprovar a existência de comportamento social no grafo e também que as rotinas exercem uma forte influência na definição do tempo de viagem e nas trajetórias.

### 4.3. Métricas Microscópicas

#### **Grau do Vértice:**

Quando avaliamos o grau do vértice numa rede veicular, é importante lembrar que durante o dia a densidade de tráfego sofre variações e conseqüentemente a quantidade de encontros que um veículo faz é influenciada pela quantidade de veículos em circulação. O gráfico na Figura 5 representa a evolução do grau dos vértices para a base de dados de Zurich. Nota-se na Figura 5-(a), que o valor do grau acompanha a ocorrência dos horários de pico, e observa-se também um aumento para o grau médio, mediano e máximo. Durante os horários de pico, como às 7am., é o momento de maior valor para o grau 1.355, o que representa a existência de veículos que fazem viagens em regiões com alta densidade de tráfego e que fazem viagens maiores. O mesmo comportamento pode ser observado durante às 3pm. com altos valores de grau. Contudo, o valor do grau mediano também apresenta um pequeno aumento, acompanhando o aumento de tráfego, com valores de 200 para o grau em horários de pico. Estes são valores aceitáveis, pois com o aumento de tráfego é esperado a ocorrência de mais encontros.

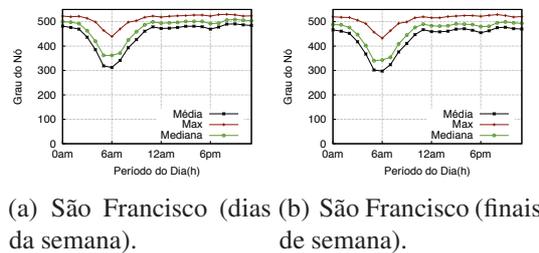
Muitas redes sociais são consideradas como livres de escala, e uma das características destas redes é que a sua distribuição de graus segue uma lei de potência [Newman 2005], poucos vértices com grau alto e muitos vértices com grau baixo. Nas Figuras 5-(b) e 5(c) pode-se observar essa distribuição para os dois horários com maior fluxo de veículos às 7am. e 3pm. em escala logarítmica. Nestes horários, encontra-se



**Figura 5. Evolução do grau do vértice durante o dia.**

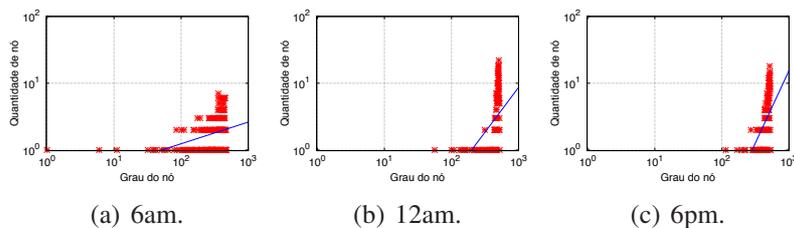
em circulação aproximadamente 100.000 veículos e as curvas nos gráficos mostram um comportamento bem próximo de uma lei de potência; poucos veículos com alto grau, e uma maior quantidade de veículos com valores menores de grau, próximo a  $10^2$ .

Para a base de dados de São Francisco, os gráficos nas Figuras 6-(a) e 6-(b), apresentam os valores pra dias da semana e finais de semana respectivamente. Nos períodos entre os horários de 5am - 7am, observa-se uma redução do grau dos vértices devido a quantidade de veículos transitando e a redução dos encontros. Além disso, percebe-se em conformidade com a densidade, que o valor médio de grau é bem elevado, levando a entender que nos horários de grande tráfego, praticamente todos os veículos se encontram. Quando analisa o comportamento durante o dia, percebe-se também uma tendência constante para esse valor, sempre acima de 450 encontros médios por veículo.



**Figura 6. Evolução do Grau do Vértice durante o dia.**

A Figura 7 apresenta a distribuição de graus para três momentos do dia, às 6am, 12am e 6pm (Fig. 7-(a), 7-(b) e 7-(c)). Pode-se notar que, com o decorrer do dia, os veículos se encontram com maior frequência. Principalmente no horário de 6pm, quando têm-se todos os veículos com graus entre  $10^2$  e  $10^3$ . Além disso, esses não é possível constatar um comportamento leis de potência para esse padrão de mobilidade.

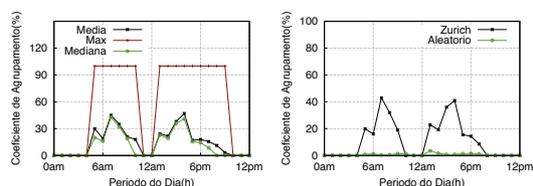


**Figura 7. Evolução do Grau do Vértice para diferentes horários.**

### **Coefficiente de Agrupamento:**

O coeficiente de agrupamento é uma métrica muito importante para análise numa rede social, pois com essa métrica é possível mensurar a organização do grafo e mensurar o quão próximo os vizinhos estão de um veículo. No contexto de uma rede veicular, o coeficiente de agrupamento pode representar os veículos na mesma direção, que tem interesses em comum ou que passam pelo mesmo lugar com frequência, isto é, um grupo de veículos que encontram em algum ponto da trajetória. A Figura 8 apresenta o valor para o coeficiente de agrupamento máximo, médio e mediano e a sua devida comparação com os grafos aleatórios durante o dia. Nota-se na Figura 8-(a) que o valor máximo do coeficiente de agrupamento é 100% durante todo o tempo dos horários de tráfego intenso, o que representa que neste período existem veículos que possuem todos os seus vizinhos conectados entre si.

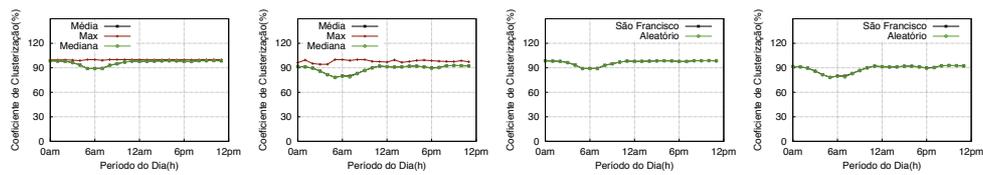
Quando existem muitos veículos em circulação na cidade, principalmente nas áreas centrais, pode aparecer muitos pontos de engarrafamento, principalmente em avenidas as quais recebem fluxos de várias ruas adjacentes. Nestas regiões, existe uma tendência forte de formação de agrupamentos de veículos, que compartilham as mesmas rotas e condições de tráfego. Ainda na Figura 8-(a) pode-se observar que o coeficiente de agrupamento médio é 50%, ou seja, metade dos vizinhos de um veículo tiveram encontro em algum momento durante sua trajetória. Quando se compara com grafos aleatórios, é possível perceber que a base de dados de Zurich apresenta um coeficiente uma ordem de grandeza maior do que os grafos aleatórios, o que nos leva a constatar que essa base apresenta fortes indicativos de comportamentos sociais, com a formação de comunidades [Watts and Strogatz 1998].



(a) Valor médio, máximo e mediano. (b) Comparação com grafos aleatórios.

**Figura 8. Evolução do Coeficiente de Agrupamento para a base de Zurich.**

A Figura 9 apresenta a evolução para o coeficiente de agrupamento durante o dia. É importante lembrar que esta rede apresenta uma alta densidade, o que já é um indicativo para formação de grupos. Nas Figuras 9-(a) e 9-(b), pode-se avaliar o valor máximo, médio e mediano para o coeficiente de agrupamento. Observa-se que o coeficiente tem o seu valor bem próximo de 90% durante todo o dia. Desta forma, praticamente todos os vizinhos de um vértice no grafo possuem conexão entre si; o que torna a rede bastante agrupada. Comparando os valores dos dias da semana para finais de semana, percebe-se uma redução do coeficiente nos finais de semana (cerca de 5%). Essa diferença pode ser atribuída ao fato de que, nos finais de semana os destinos são variados, o que implica numa variação das trajetórias efetuadas pelos veículos. Ao comparar os resultados com grafos aleatórios (Fig. 9-(c) e 9-(d)), é possível verificar praticamente a existência de um único agrupamento no grafo.

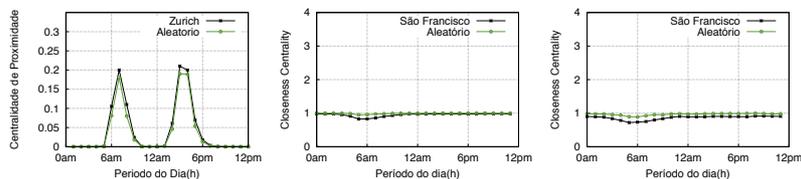


(a) Dias da semana. (b) Finais de semana. (c) Dias da semana. (d) Finais de semana.

**Figura 9. Evolução do Coeficiente de Agrupamento para a base de São Francisco.**

### Centralidade de Proximidade:

A evolução da centralidade de proximidade (*Closeness Centrality*) no grafo é importante para entender o quão próximo os vértices se encontram. A Figura 10 apresenta o valor médio para a *closeness centrality* durante o dia para as bases de Zurich e São Francisco. A Figura 10-(a) apresenta os valores para a base de dados de Zurich. Pode-se verificar um crescimento no valor da centralidade de proximidade nos horários de pico. Por definição, quanto maior for o valor desse coeficiente, menor será a distância entre veículos. Assim, quando se tem muitos veículos, a tendência é que os mesmos se encontrem com maior frequência, principalmente nos horários de pico, nos quais a probabilidade de engarrafamentos é maior. Para a base de São Francisco (Figura 10-(b) e 10-(c)) pode-se observar um comportamento semelhante. Nos horários de pouco tráfego o valor do coeficiente sofre um pequeno decréscimo. Durante o restante do dia, devido à alta densidade de arestas, o valor desse coeficiente tende a ficar constante. Para ambas as bases de dados a comparação com os grafos aleatórios não gerou variações expressivas.



(a) Zurich. (b) São Francisco (dias da semana). (c) São Francisco (finais de semana).

**Figura 10. Evolução da Centralidade de Proximidade durante o dia.**

## 5. Conclusões e Trabalhos Futuros

Neste trabalho foi apresentada uma análise social para duas bases de dados conhecidas na literatura, a base de Zurich e a base de São Francisco. São duas bases que possuem comportamentos e rotinas diferentes. A primeira representa o tráfego de veículos pessoais numa região de grande cidade e a segunda base um conjunto de táxis se deslocando para atender passageiros, também numa grande cidade. A análise da base de dados de Zurich apontou fortes características sociais, como a presença de comunidades comprovada pela métrica coeficiente de agrupamento e as distâncias curtas entre os vértices no grafo. Pode-se comprovar também a existência de interesses comuns e rotinas nesta base. Entretanto, a base de São Francisco, por apresentar movimentos de veículos de mesma natureza (táxis), em regiões próximas, apresentou uma grande densidade de encontros, com a estrutura bem próxima de um grafo completo. Constata-se também nessa base que a diversidade

de destinos de um táxi e o seu constante deslocamento diário, deixa-o mais suscetível a encontrar outros veículos. Essas características reforçam a ideia que a análise de um trace sob a perspectiva social deve considerar as peculiaridades da cidade, dos tipos de veículos e o seu padrão de mobilidade. Contudo, ao analisar rotinas e padrões de mobilidade de veículos, é importante ressaltar que a existência de similaridade nas rotas deve considerar toda a rotina de um condutor e a sua interação com os demais, durante todo o seu trajeto.

Como trabalhos futuros, pretende-se analisar outras bases, além de caracterizar essa mobilidade em função de tempo e espaço, tentando quantificar diferenças entre as métricas por regiões numa cidade. Além disso, almeja-se o refinamento desta análise, com a variação dos parâmetros utilizados na construção dos grafos e avaliação de outras métricas. Outro ponto importante a investigar é a aplicação dessas métricas para melhorias de protocolos em redes veiculares. Protocolos de disseminação de dados podem fazer uso de métricas sociais para alcançar melhorias na cobertura e redução do número de transmissão. Além disso, métricas que quantificam interesses comuns podem ser úteis na disseminação de propagandas e informativos na rede. Desta forma, pretende-se aplicar conceitos sociais no desenvolvimento de protocolos de roteamento em redes veiculares.

## Referências

- Boukerche, A., Oliveira, H. A. B. F., Nakamura, E. F., and Loureiro, A. A. F. (2008). Vehicular ad hoc networks: A new challenge for localization-based systems. *Computer Communications*, 31(12):2838–2849.
- Cunha, F., Carneiro Viana, A., Mini, R. A. F., and A.F. Loureiro, A. (2013). How effective is to look at a vehicular network under a social perception? In *1st International Workshop on Internet of Things Communications and Technologies (IoT'13) (IoT'2013)*, Lyon, France.
- Faezipour, M., Nourani, M., Saeed, A., and Addepalli, S. (2012). Progress and challenges in intelligent vehicle area networks. *Communications ACM*, 55(2):90–100.
- Fiore, M. and Häiri, J. (2008). The networking shape of vehicular mobility. In *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing, MobiHoc '08*, pages 261–272, New York, NY, USA. ACM.
- Hossmann, T., Legendre, F., and Spyropoulos, T. (2009). From Contacts to Graphs: Pitfalls in Using Complex Network Analysis for DTN Routing. In *INFOCOM Workshops 2009, IEEE*, pages 1–6. IEEE.
- Johnson, N. L. and Kotz, S. (1977). *Urn Models and Their Applications: An Approach to Modern Discrete Probability Theory*. Wiley, New York.
- Liu, X., Li, Z., Li, W., Lu, S., Wang, X., and Chen, D. (2012). Exploring social properties in vehicular ad hoc networks. In *Proceedings of the Fourth Asia-Pacific Symposium on Internetware, Internetware '12*, pages 24:1–24:7, New York, NY, USA. ACM.
- Loulloudes, N., Pallis, G., and Dikaiakos, M. D. (2010). The dynamics of vehicular networks in urban environments. *CoRR*, abs/1007.4106.
- Naumov, V., Baumann, R., and Gross, T. (2006). An evaluation of inter-vehicle ad hoc networks based on realistic vehicular traces. In *Proceedings of the 7th ACM international symposium on Mobile ad hoc networking and computing, MobiHoc '06*, pages 108–119, New York, NY, USA. ACM.
- Newman, M. E. J. (2005). Power laws, pareto distributions and zipf's law. *Contemporary Physics*, 46:323–351.
- Piorowski, M., Sarafijanovic-Djukic, N., and Grossglauser, M. (2009). Crawdad trace. Downloaded from <http://crawdad.cs.dartmouth.edu/>.
- Schoch, E., Kargl, F., Weber, M., and Leinmuller, T. (2008). Communication patterns in vanets. *Communications Magazine, IEEE*, 46(11):119–125.
- Scott, J. (2000). *Social Network Analysis: A Handbook*. Sage Publications, second. edition.
- Tostes, A. I. J., de L. P. Duarte-Figueiredo, F., Assunção, R., Salles, J., and Loureiro, A. A. F. (2013). From data to knowledge: City-wide traffic flows analysis and prediction using bing maps. In *SIGKDD International Workshop on Urban Computing, Urb-Comp'13*, Chicago, USA. ACM.
- Uppoor, S. and Fiore, M. (2012). Insights on metropolitan-scale vehicular mobility from a networking perspective. In *Proceedings of the 4th ACM international workshop on Hot topics in planet-scale measurement, HotPlanet '12*, pages 39–44, New York, NY, USA. ACM.
- Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442.
- Wellman, B. (2001). Computer networks as social networks. *Science*, 293(5537):2031–2034.